Niveau(x) d'enseignement: 0 Numéro EDUQ : 8470678
Centre d'intérêt: 540

Scallon, Gérard, 1938-.

L'évaluation formative, un outil au service de l'enseignement et de l'apprentissage / par Gérard Scallon. -- [Québec] : Université Laval, Département de mesure et évaluation, 1980.
39 f. ; 28 cm.

Notes bibliogr. au bas des p.

Descripteur(s): *Evaluation formative *Mesure *Instrument d'évaluation

L'évaluation formative est une mesure ou une observation qui intervient de façon continue dans la progression de l'étudiant engagé dans une démarche d'apprentissage. Le présent document porte sur divers aspects de cette évaluation. Il traite, d'abord, de la fonction diagnostique de l'évaluation formative. Il présente, ensuite, certaines méthodes proposées pour décrire la structure des objectifs d'un cours, dans une perspective de planification des instruments de l'évaluation formative. La troisième partie décrit les instruments pouvant être utilisés: a) test centré sur un ou plusieurs objectifs; b) mesure reliée à un domaine; c) analyse et mesure des concepts; d) détection des erreurs caractéristiques; e) épreuve hiérarchique. La dernière partie traite de l'utilisation des résultats.

Type de document: r

EDUQ 8470678

L'EVALUATION FORMATIVE:

un outil au service de

l'enseignement et de l'apprentissage

par Gérard SCALLON, professeur titulaire,
Département de Mesure et Evaluation,
Faculté des sciences de l'éducation,
Université Laval.

JANVIER 1980

TABLE DES MATIÈRES

(page)

L'ÉVALUATION FORMATIVE ET LE DIAGNOSTIC PÉDAGOGIQUE

définition de l'évaluation formative. 1

la fonction diagnostique de l'évaluation formative: le diagnostic pédagogique 1

le diagnostic pédagogique et l'évaluation diagnostique 5

exigences de l'évaluation formative 6

LE MODE D'ORGANISATION DES OBJECTIFS 9

LES INSTRUMENTS DE L'ÉVALUATION FORMATIVE

le test centré sur un ou plusieurs objectifs . . . 18

la mesure reliée à un domaine 20

l'analyse et la mesure des concepts: le modèle de Markle et Tiemann 26

La détection des erreurs caractéristiques: la rédaction systématique des leurres pour des questions à choix de réponse. 30

l'épreuve hiérarchique. 31

L'UTILISATION DES RÉSULTATS EN ÉVALUATION FORMATIVE: la feuille de route individuelle et la matrice "étudiants x objectifs". 35

L'EVALUATION FORMATIVE ET LE DIAGNOSTIC PEDAGOGIQUE

Définition de l'évaluation formative.-

Dans les écrits les plus récents qui ont tenté de clarifier et d'expliciter les diverses fonctions de l'évaluation des étudiants, on trouve quatre types d'évaluation: de classement, diagnostique, formative et sommative (1). De ces quatre types d'évaluation, l'évaluation formative est sans contredit le type qui est le plus près du processus d'enseignement et d'apprentissage. Alors que l'évaluation diagnostique et l'évaluation de classement surviennent habituellement avant une démarche d'apprentissage et que l'évaluation sommative survient après, l'évaluation formative se réalise pendant le déroulement même de cette démarche. On peut définir l'évaluation formative des étudiants comme étant:

> un ensemble ou une collection d'instruments de mesure ou d'observation conçus spécifiquement dans le but d'intervenir de façon continue dans la progression de chaque étudiant engagé dans une démarche d'apprentissage.

La figure 1 (page 2) présente une démarche théorique d'apprentissage à l'intérieur d'une étape: la poursuite d'un objectif terminal et de trois objectifs intermédiaires devant faciliter, à titre de jalons, l'atteinte de cet objectif terminal. Chaque losange de la figure 1 symbolise une décision pouvant modifier, s'il y a lieu, le cheminement ou la progression de chaque étudiant à travers l'ensemble des quatre objectifs. La décision de faire progresser ou non l'étudiant vers un autre objectif doit être éclairée par un instrument de mesure approprié. L'ensemble des quatre épreuves de rendement qui correspondraient aux losanges de la figure 1 définirait l'évaluation formative des étudiants pour l'étape considérée.

La fonction diagnostique de l'évaluation formative: le diagnostic pédagogique.-

Pour chaque prise de décision réalisée selon l'esprit de l'évaluation formative il nous faut considérer les prescriptions d'ordre pédagogique qui s'ensuivent selon que la décision est positive ou négative:

1- décision positive: si l'étudiant a atteint ou maîtrisé un objectif il est amené à aborder et à poursuivre un nouvel objectif dans le cadre de sa démarche d'apprentissage;

2- décision négative: si l'étudiant ne maîtrise pas suffisamment l'objectif son cheminement est alors momentanément interrompu. Le problème qui se pose alors est de prescrire un enseignement correctif approprié.

1- Peter W. Airasian et George F. Madaus, "Functional Types of Student Evaluation", Measurement and Evaluation in Guidance, Vol. 4, No. 4, 1972, pp. 221-233.

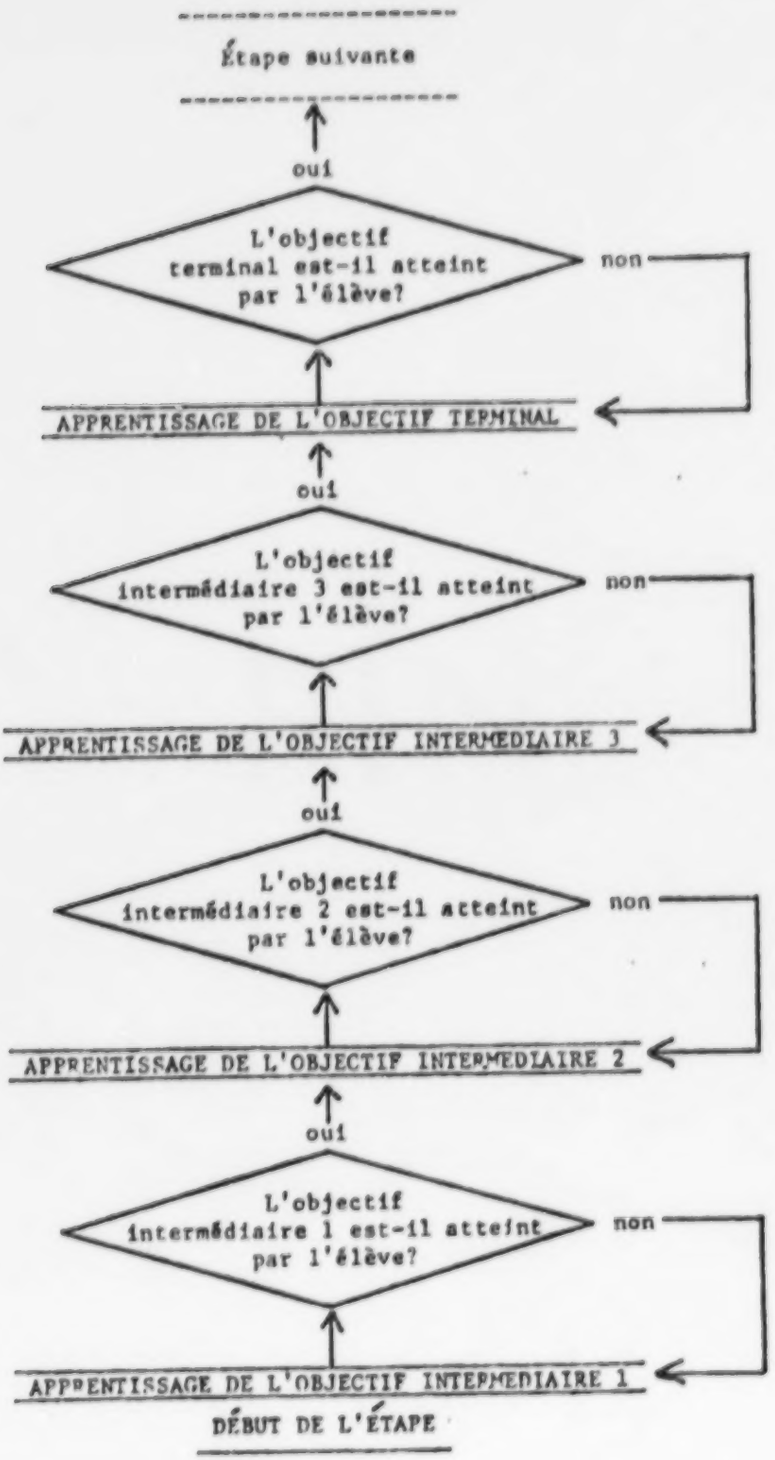


Figure 1: Evaluation formative du progrès de chaque étudiant dans une démarche d'enseignement et d'apprentissage comportant trois objectifs intermédiaires et un objectif terminal.

Dans le cas d'une décision négative, l'enseignement correctif peut se faire selon divers dispositifs que l'on peut regrouper en trois catégories:

a) soit que l'étudiant reprenne les activités ou exercices qui ont servi à l'apprentissage initial,

b) soit que l'étudiant ne reprenne seulement qu'une partie des activités ou exercices ayant servi à l'apprentissage initial et ce d'après des difficultés particulières décelées chez l'étudiant,

c) soit que l'étudiant reprenne des activités ou exercices non directement reliés à l'objectif non maîtrisé mais en rapport avec des objectifs intermédiaires (ou pré-requis à l'objectif non maîtrisé).

Le diagramme qui permet de visualiser la place de ces catégories d'enseignement correctif au niveau de chaque décision en évaluation formative apparaît à la figure 2 (page 4).

Le choix du dispositif le plus adéquat pour réaliser l'enseignement correctif dépend de plusieurs facteurs dont la nature même de l'objectif qui n'a pas été maîtrisé (domaine --cognitif --affectif --psychomoteur; niveau hiérarchique de l'objectif dans une taxonomie), de la matière, du degré de linéarité (ou du caractère cyclique) de la démarche d'apprentissage, du mode d'enseignement et des activités d'apprentissage utilisées. De plus, les deux dernières catégories de dispositif pour l'enseignement correctif sont plus individualisées que la première et leur degré de sophistication devrait dépendre de l'importance qui doit être accordée à l'atteinte de l'objectif.

Il est important de préciser tout de suite que la prescription d'un enseignement correctif qui va plus loin que la reprise intégrale de l'apprentissage initial exige que l'on ait recours à des instruments de mesure de nature un peu spéciale, ce dont il sera question plus loin dans ce texte.

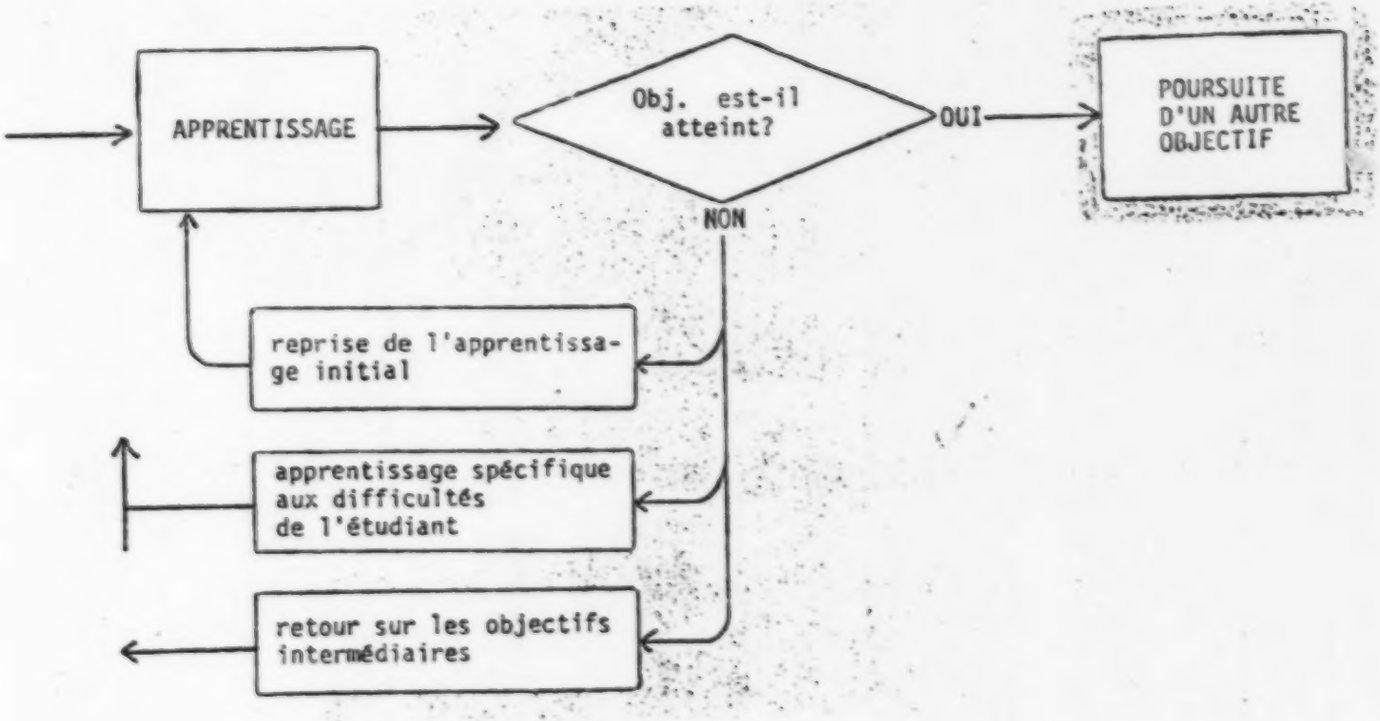


Figure 2: catégories d'enseignement correctif pouvant être impliquées au niveau d'une décision particulière en évaluation formative.

Le diagnostic pédagogique et l'évaluation diagnostique.-

Le diagnostic pédagogique est une fonction essentielle de l'évaluation formative (1)(2). Toutefois il ne faudrait pas confondre cette fonction avec celle de l'évaluation diagnostique (type d'évaluation distinct de l'évaluation formative). L'évaluation diagnostique a trait à de nombreux facteurs pouvant expliquer des difficultés d'apprentissage (intérêts, attitudes, motivation, situation familiale et socio-économique, aptitudes etc.) alors que les instruments de l'évaluation formative ont pour objet le rendement scolaire proprement dit. Loin de s'opposer, ces deux types d'évaluation peuvent se compléter dans la détection des causes pouvant expliquer des difficultés d'apprentissage.

Pour bien saisir la relation qui peut exister entre l'évaluation formative (diagnostic pédagogique) et l'évaluation diagnostique il suffit de mettre en évidence le type d'information qu'il est nécessaire de recueillir, en présence d'un échec dans l'atteinte d'un objectif par un étudiant, afin de prescrire un correctif approprié qui, dans certains cas, ne sera pas forcément de nature pédagogique.

La non maîtrise d'un objectif par un étudiant doit être interprétée selon deux dimensions:

a) les réussites et les échecs antérieurs manifestés par cet étudiant dans le domaine des habiletés dont l'objectif fait partie,

b) le nombre d'étudiants, dans le groupe d'apprentissage dont l'étudiant fait partie, qui ont échoué (ou réussi) dans l'atteinte de l'objectif d'intérêt.

L'usage que l'on peut faire de ces deux dimensions ou perspectives dans l'interprétation d'un échec apparaît assez clairement dans ce texte d'Anna Bonboir:

> "L'échec se présente, de son côté, soit comme une déficience généralisée, soit comme un état particulier: déficience généralisée chez un sujet échouant dans l'ensemble de ses apprentissages, sauf peut-être dans l'un d'entre eux avec lequel il entretient une sorte de connivence qu'il faudrait expliquer; déficience généralisée, s'éten-

---

1- P.W. Airasian et G.F. Madaus, op. cit., pp. 224-225;

2- Benjamin S. Bloom, J. Thomas Hastings et George F. Madaus, "Formative Evaluation", Chap. 6 dans: Handbook on Formative and Summative Evaluation of Student Learning, McGraw-Hill Book Co., Toronto, 1971, pp. 133-135;

> dant à l'ensemble des apprentissages chez tous les sujets du groupe étudié ou de la classe à laquelle le test diagnostique fut appliqué; déficience spécifique bien localisée et bien décrite, commune à tout le groupe étudié, sauf peut-être quelques exceptions; enfin, déficience spécifique particulière à un sujet ou à quelques uns éprouvant des difficultés propres dans un secteur déterminé.
>
> Suivant qu'il s'agit, dans la manifestation de l'échec, de l'une ou l'autre de ces éventualités, les déductions hypothétiques relatives à l'étiologie du mal se diversifient.(1)

La démarche diagnostique déclenchée par un échec et devant être suivie d'une prescription dans l'ordre des correctifs à apporter peut être opérationnalisée au moyen d'un ordinogramme tel que le fait voir la figure 3 (page 7). Dans cet ordinogramme les réponses aux questions posées en succession nous amènent à prescrire des correctifs assez variés:

1.- le correctif de nature pédagogique qui correspond à l'enseignement correctif proprement dit (l'étudiant doit reprendre l'apprentissage initial selon un des dispositifs décrits à la figure 2);

2.- l'amélioration de la stratégie d'enseignement et d'apprentissage suivie d'une reprise de l'apprentissage par l'ensemble des étudiants qui ont éprouvé des difficultés;

3.- un sondage plus approfondi des facteurs pouvant expliquer une déficience persistante dans une matière donnée au niveau des intérêts, de la motivation, des aptitudes, de la famille, de l'état de santé de l'étudiant etc. C'est dans le cadre de ce troisième type de prescription que L'EVALUATION DIAGNOSTIQUE (comme type d'évaluation) peut être invoquée.

Exigences de l'évaluation formative.-

En résumé, l'évaluation formative est par définition un type d'évaluation centré sur un ensemble de décisions pouvant affecter la progression de chaque étudiant dans une démarche d'apprentissage. Chacune de ces décisions peut conduire à un correctif pédagogique qui implique exclusivement l'étudiant selon divers dispositifs plus ou moins sophistiqués. Enfin, en évaluation formative, il faut être à l'affût de diverses éventualités dans l'interprétation d'un échec pour que, le cas échéant, des correctifs n'impliquant pas exclusivement l'étudiant ou des correctifs autres que pédagogiques puissent être apportés.

1.- Anna Bonboir, La pédagogie corrective, Presses Universitaires de France, Paris, 1970, p. 60.

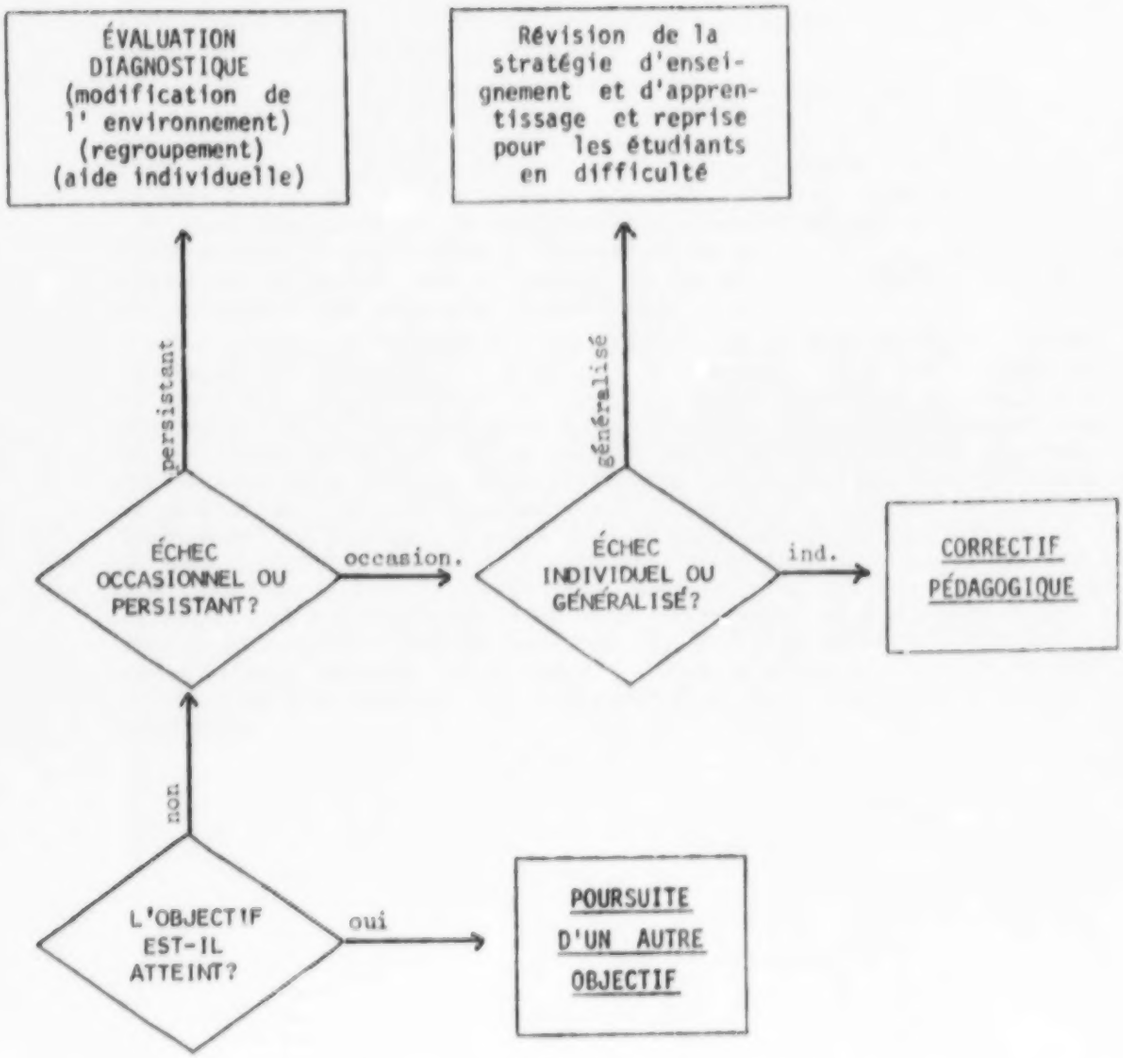


Figure 3: Ordinogramme décrivant le processus de diagnostic dans le cadre de l'évaluation formative de l'apprentissage.

L'analyse décisionnelle des interventions qui peuvent être faites dans le cheminement d'un étudiant suppose au départ que l'on a clairement identifié des objectifs pédagogiques à poursuivre. De plus, ces objectifs doivent être reliés entre eux (logiquement) de manière à définir un itinéraire, un cheminement ou un parcours. C'est d'ailleurs l'existence d'un tel itinéraire qui justifie le recours à un type d'évaluation comme l'évaluation formative.

Au niveau de chaque objectif, ou de chaque jalon de l'itinéraire, la décision d'interrompre la progression de l'étudiant, si elle doit être prise, doit être assortie de prescriptions qui viennent corriger des lacunes particulières aux difficultés éprouvées par l'étudiant. Le test centré sur un ou plusieurs objectifs sera dans certains cas insuffisant pour prescrire un enseignement correctif adéquat et il faudra alors recourir à des modèles d'instruments de nature un peu plus sophistiquée qui demeurent dans le domaine des tests dits "critériés" mais qui ont une portée beaucoup plus diagnostique.

Enfin, pour ce qui est des éventualités pouvant expliquer l'échec, le diagnostic pédagogique, assumé dans le cadre de l'évaluation formative, ne saurait se restreindre à l'observation isolée d'un échec manifesté par un étudiant dans l'atteinte d'un objectif. D'une part, le cheminement passé de l'étudiant permettrait de dire si cet échec s'ajoute à de nombreux échecs antérieurs (signal d'une difficulté persistante ou d'une déficience généralisée chez cet étudiant) ou s'il s'agit simplement d'un événement occasionnel ou passager. D'autre part le rendement d'autres étudiants à un même objectif permettrait de dire si l'échec observé est un événement individuel ou collectif (généralisé). Dans les deux cas il s'agit bien sûr de déductions hypothétiques. Pour interpréter ainsi un échec dans cette double perspective l'évaluation formative doit être assortie d'une certaine méthode comptable ou de tenue de livre au moyen de laquelle la feuille de route d'un étudiant est juxtaposée à la feuille de route d'autres étudiants.

LE MODE D'ORGANISATION DES OBJECTIFS

Dans cette section nous allons présenter certaines méthodes qui ont été proposées pour décrire la structure des objectifs d'un cours, d'une partie de cours ou d'une unité d'enseignement et d'apprentissage. Ces méthodes seront examinées dans une perspective de planification des instruments de l'évaluation formative. Afin de ne pas perdre de vue toute la complexité de l'évaluation du rendement scolaire nous allons démontrer qu'il est possible de planifier les instruments de l'évaluation formative et ceux de l'évaluation sommative dans une même démarche.

Bloom, Hastings et Madaus ont proposé un schéma d'analyse(1) d'un cours ou d'un programme d'études selon deux dimensions: le contenu et les comportements visés (ou classes de comportements). L'analyse de contenu ne présente pas de difficulté particulière puisque la démarche nous est familière. L'opération consiste essentiellement à subdiviser la matière (d'un cours par exemple) en éléments, notions ou sous-notions. A chaque élément de contenu sont rattachés un ou plusieurs verbes d'action empruntés à une taxonomie comme celle de Bloom et collaborateurs (par exemple). Le résultat de ce mode d'analyse de la matière et des objectifs d'un cours est représenté par un tableau de spécirications, tableau à double entrée puisqu'il y a deux dimensions considérées. Dans ce tableau, l'intersection d'un élément de contenu avec une catégorie ou niveau taxonomique peut donner lieu à un objectif spécifique qui deviendra un jalon dans une progression. Le type de tableau de spécifications proposé par Bloom, Hastings et Madaus pour l'évaluation formative ressemble de très près à celui qui est habituellement utilisé pour planifier des examens dans le cadre de l'évaluation sommative. La marque distinctive du tableau de spécifications proposé pour l'évaluation formative réside dans les liens qui apparaissent entre certaines intersections de façon à préciser un cheminement ou un itinéraire particulier. La figure 4 (page 10) présente un tableau de spécifications ainsi conçu avec une unité d'un cours de statistique descriptive traitant des mesures de tendance centrale et de dispersion. Dans la colonne "éléments de contenu" on voit apparaître les notions et sous-notions regroupées sous deux thèmes majeurs: "mesures de tendance centrale" et "mesures de dispersion". C'est le résultat de la structuration du contenu pour cette unité d'intérêt. En horizontale (titres des colonnes) apparaissent les niveaux taxonomiques pertinents avec la matière et les buts du cours de statistique descriptive, depuis la connaissance des termes jusqu'à l'application. Le niveau intermédiaire, "habileté à utiliser des processus et des procédés", est récent dans la préoccupation de Bloom et n'apparaît pas comme tel dans la taxonomie des objectifs pédagogiques(2).

La figure 4 fait voir les relations entre certaines intersections et met ainsi en relief le cheminement (présumément idéal) à sui-

---

1- B.S. Bloom, J.T. Hastings et G.F. Madaus, op. cit., Chap. 6;

2- ibid., p. 119.

ELEMENTS DE CONTENU	connaissance de la terminologie	habileté à utiliser des processus et des procédures (calcul)	habileté à appliquer
<u>Mesures de tendance centrale</u>:			
moyenne arith.	1	7	13
médiane	2	8	13
mode	3	9	13
<u>Mesures de dispersion</u>:			
étendue simple	4	10	14
écart semi-interquartile	5	11	14
écart-type	6	12	14

Figure 4: exemple de tableau de spécifications selon le type proposé par Bloom, Hastings et Madaus (1971) pour l'évaluation formative appliquée à une unité de cours en statistique descriptive.

vre dans la poursuite des objectifs du cours. Les nombres ordinaux assignés aux intersections du tableau de spécifications indiquent également un cheminement entre les diverses sous-unités de contenu. Par exemple, le tableau de spécifications de la figure 4 indique que les six sous-notions seront abordées successivement au niveau de la connaissance de la terminologie et seront reprises ensuite au niveau du calcul. On aurait pu prévoir un cheminement différent (sans modifier l'ordonnancement des objectifs) en proposant un itinéraire continu d'un niveau d'objectif à l'autre avec les mesures de tendance centrale suivi d'un itinéraire continu avec les mesures de dispersion. Dans les deux cas la "configuration" du tableau de spécifications est la même et les liens entre les intersections sont constants. La différence réside dans le parcours général indiqué par les nombres ordinaux (qui peuvent servir également à identifier les objectifs)apparaissant aux intersections.

Le choix d'un parcours idéal repose sur la nécessité plus ou moins grande d'aborder les objectifs qui sont reliés entre eux (selon l'ordonnancement) d'une manière continue de sorte que des objectifs sont abordés immédiatement avant d'autres objectifs auxquels ils sont pré-requis.

Le tableau de spécifications de Bloom, Hastings et Madaus peut s'avérer un outil fort commode pour structurer les objectifs poursuivis dans un cours ou dans un programme d'études. Il met en relief les deux facettes d'un cheminement ou d'une progression: (a) l'ordonnancement des objectifs et (b) le parcours choisi qui respecte cet ordonnancement. Le tableau sert ainsi à déterminer quels objectifs (intersections) feront l'objet de l'évaluation formative. Toutefois l'utilisation d'un tel instrument suppose que les objectifs d'un cours peuvent être ramenés essentiellement à quelques niveaux ou éléments d'une taxonomie qui constitue en quelque sorte un schéma de classification de ces objectifs. Le problème de traduire convenablement chaque catégorie taxonomique en objectif opérationnel, et ce pour chaque subdivision du contenu, demeure entier. Le lecteur intéressé à cette question trouvera dans l'ouvrage de De Landsheere et De Landsheere (1) un bilan assez représentatif des divers essais d'opérationnalisation des objectifs de la taxonomie et un aperçu des nombreuses difficultés qui persistent encore aujourd'hui en ce domaine.

1- Gilbert De Landsheere et Viviane De Landsheere, Définir les objectifs de l'éducation, Bruxelles: éditions Georges Thone, s.a., 1975, pp. 84-94;

Le tableau de spécifications de Bloom et collaborateurs est le résultat d'une analyse globale de la matière et des objectifs d'un cours ou d'une unité imposante d'enseignement et d'apprentissage. Travailler à un niveau aussi global est une question de sens commun et il ne semble pas exister de techniques particulières que l'on puisse proposer aux concepteurs de programmes pour y arriver comme l'affirment Gagné et Briggs (1).

Dans le cadre de l'enseignement programmé des techniques assez précises ont été développées pour analyser une portion de matière ou un thème précis en notions élémentaires à être disposées par la suite dans un ordre logique ou psychologique. Au départ, certaines de ces techniques se sont limitées exclusivement à une analyse de contenu débouchant sur une ordonnance d'informations (par exemple la décomposition d'un cours en notions et sous-notions par la méthode de Le Xuan, ou par la méthode des graphes de Morganov). La nécessité de plus en plus reconnue d'expliciter les intentions pédagogiques en termes comportementaux a fait que certaines de ces méthodes ont pu être adaptées à l'analyse psychologique ou comportementale d'une matière. Ainsi, l'ouvrage de D. Leclercq et collaborateurs (2) présente un exemple d'application et d'intégration de trois méthodes d'analyse de la matière qui ont servi à l'élaboration d'un cours d'introduction au langage FORTRAN.

Dans un cadre plus théorique Robert Gagné a proposé un modèle d'analyse selon des types d'apprentissage. Ces types d'apprentissage présentent un double intérêt pour l'évaluation formative: leur caractère hiérarchique et leur résonnance académique (objectifs de comportement). La théorie de Gagné a fait l'objet de plusieurs raffinements depuis son apparition vers les années 60 et il serait onéreux d'en tracer ici un historique complet. L'ouvrage de Gagné et Briggs (3) permet toutefois de décrire les grandes lignes du modèle tel qu'il apparaissait jusqu'à tout récemment. Les auteurs présentent cinq types de "capacités" résultant de l'apprentissage:

- -les habiletés motrices
- -l'information verbale
- -les habiletés intellectuelles
- -les attitudes
- -les stratégies cognitives.

L'acquisition de chacune de ces capacités doit suivre des règles précises graduant les interventions: ce sont les conditions d'apprentissage. Les hiérarchies d'apprentissage apparaissent au niveau des habiletés intellectuelles et constituent les conditions mêmes d'apprentissage pour ce type de capacité.

1- Robert Gagné et Leslie Briggs, Principles of Instructional Design Holt Rinehart, New York, 1974, p. 100;

2- D. Leclercq, J. Donnay, R. De Bal et P. Lambrecht, Construire un cours programmé, Bruxelles: Editions Labor (Paris: Fernand Nathan), 1973;

3- Robert Gagné et Leslie Briggs, op. cit., p. 105.

Gagné et Briggs définissent ainsi une hiérarchie d'apprentissage:

> The learning hierarchy is an arrangement of intellectual skill objectives into a pattern which shows the prerequisite relationships among them. Beginning with a particular objective (often a lesson objective), the learning hierarchy shows which intellectual skills are prerequisite; having identified this second set of skills, the prerequisites of each of these is in turn indicated, and this process continues until one has displayed in a bottom "row" the most elementary intellectual skills with which one needs to be concerned. (1)

Ainsi définie, la hiérarchie d'apprentissages présente une analogie étroite avec la notion de réseau d'objectifs ou d'itinéraire dont il a été question précédemment. Elle constitue donc essentiellement un cheminement, un itinéraire à parcourir, parmi des capacités subordonnées les unes aux autres et convergeant vers la maîtrise d'une tâche terminale.

Pour le practicien de l'évaluation ou le concepteur de programmes, l'élaboration d'une hiérarchie d'apprentissages n'est pas une entreprise aisée. Les nombreux textes et ouvrages reliés à la théorie de Gagné contiennent cependant une multitude d'exemples de hiérarchies d'apprentissages. Depuis l'origine et cette approche, le principe d'analyse devant servir à l'élaboration de hiérarchies se ramène à une question, d'abord posée à l'endroit de l'objectif terminal, et ensuite posée à l'endroit de chacun des objectifs intermédiaires considérés en succession. Devant un objectif terminal que l'on prend pour cible la question posée est: qu'est-ce que l'individu (l'étudiant) doit savoir faire pour réussir cette (nouvelle) tâche si on lui donne uniquement des instructions verbales (Gagné et Paradise, 1961, p. 2).

Pour le concepteur de programmes la connaissance des types d'apprentissage proposés dans la théorie de Gagné est un atout précieux susceptible de lui faciliter la tâche dans l'application de ce principe d'analyse.

La figure 5 (page 14) présente un exemple de hiérarchie d'apprentissages ayant pour objectif terminal la capacité d'accorder le participe passé conjugué avec l'auxiliaire 'avoir' dans des phrases simples. La hiérarchie a été obtenue en appliquant le principe énoncé auparavant sans se préoccuper de la correspondance entre les objectifs intermédiaires et les types d'apprentissage de la théorie de Gagné.

Comme le tableau de spécifications de Bloom et collaborateurs, la hiérarchie d'apprentissage présentée à la figure 5 indique les deux facettes du cheminement proposé à chaque étudiant: (a) l'ordonnancement des objectifs intermédiaires en vue de l'objectif terminal et (b) le parcours décrit par les nombres ordinaux tout en respectant cet ordonnancement. L'avantage indéniable d'un tel outil est sans contredit

---

1- Robert Gagné et Leslie Briggs, op. cit., p. 109.

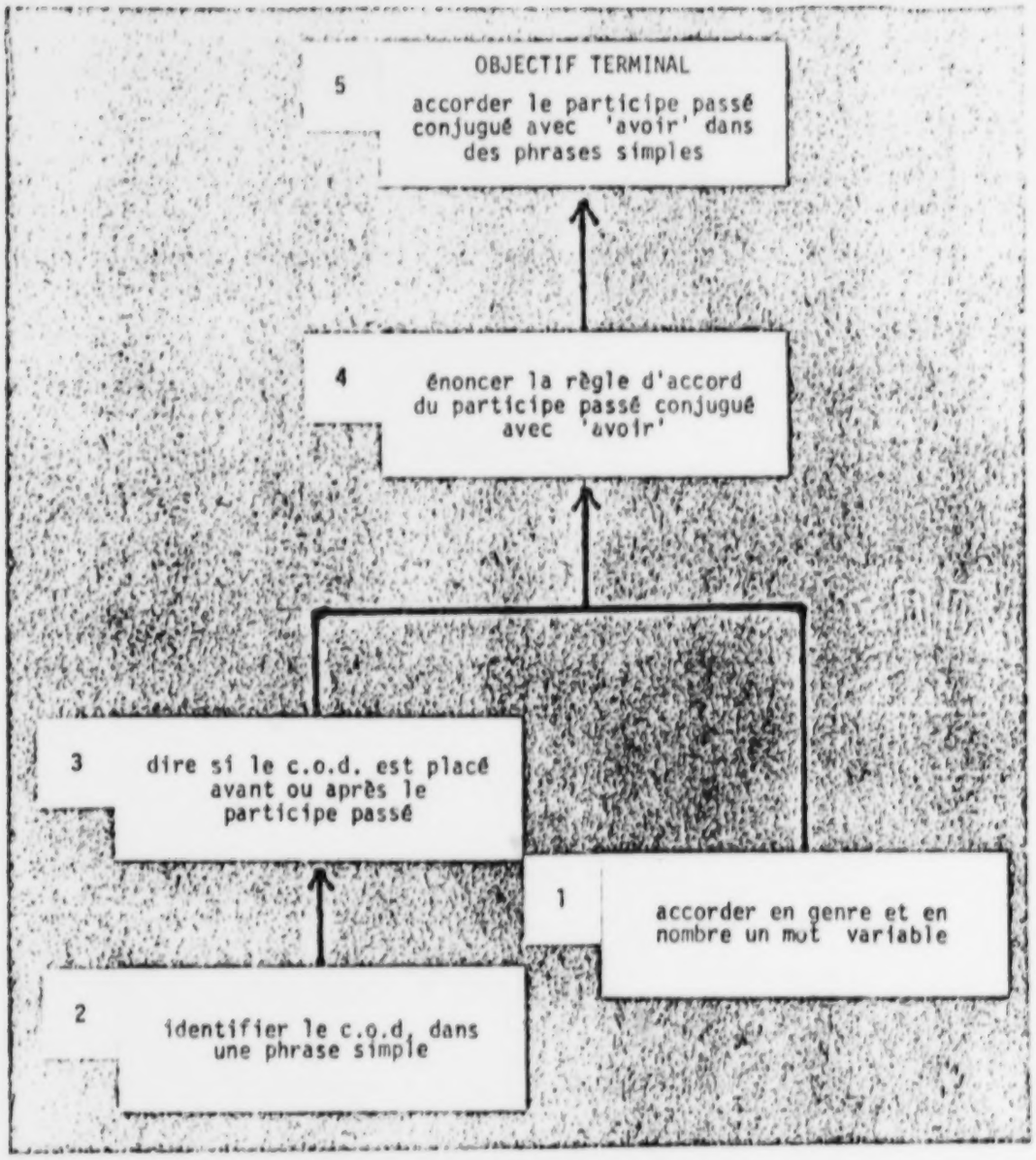


Figure 5: exemple de hiérarchie d'apprentissage menant à l'accord du participe passé conjugué avec l'auxiliaire 'avoir'.

le fait que les jalons de la hiérarchie sont précisés au moyen d'objectifs de comportement ce qui serait difficilement concevable pour un tableau de spécirications se rapportant à un cours ou à une unité importante d'enseignement.

Une hiérarchie comme celle qui est illustrée par la figure 5 doit être utilisée avec réserve. D'ailleurs, pour tout procédé d'ordonnancement des objectifs le résultat obtenu demeure toujours hypothétique et les liens hiérarchiques indiqués par l'ordonnancement doivent être démontrés expérimentalement. De plus il n'est pas assuré que le cheminement proposé par l'ordonnancement et par le parcours (nombres ordinaux),dans une hiérarchie d'objectifs, conviennent à tous les étudiants.

Jusqu'ici nous avons présenté deux méthodes pour décrire la structure des objectifs d'un cours ou d'une unité relativement restreinte d'enseignement et d'apprentissage avec la préoccupation d'établir un itinéraire de base pouvant justifier le recours à l'évaluation formative. Il nous rests à commenter brièvement de quelle façon on peut se servir de ces méthodes pour préciser quels objectifs feront l'objet de l'évaluation formative et quels objectifs feront l'objet de l'évaluation sommative.

Classiquement la question ne semble s'être jamais posée là où l'évaluation dite sommative occupait presqu'exclusivement le champ de l'évaluation du rendement scolaire. D'ailleurs les examens terminaux ou de fin d'année devaient couvrir, au moyen d'un échantillon de questions, l'ensemble des objectifs d'un cours ou d'un programme d'études. Si avec l'avènement de l'évaluation formative on devait de plus soumettre tous les objectifs, de pareille façon, à ce nouveau type d'évaluation, celà entraînerait inévitablement une redondance des plus déplorables en matière d'évaluation.

La structuration des objectifs a pour conséquence de mettre en relief des objectifs qui sont plus importants que d'autres et qui sont habituellement le terme d'une unité d'enseignement et d'apprentissage à l'intérieur d'un cours ou d'un programme d'études. Pour atteindre ces objectifs ultimes ou terminaux on fera poursuivre à l'étudiant des objectifs qui leur sont intermédiaires. La raison d'être de ces objectifs dits intermédiaires est de faciliter l'atteinte d'un ou de plusieurs objectifs terminaux. Enfin il peut arriver que l'on identifie des objectifs isolés qui appartiennent à la catégorie "nice to know" et qui font office de "remplissage". De tels objectifs ont généralement une importance très discutable sur le plan de l'évaluation.

Compte tenu des fonctions de l'évaluation formative et celles de l'évaluation sommative on peut avancer quelques principes pouvant guider notre réflexion dans le partage des objectifs dont on veut évaluer l'atteinte.

L'évaluation sommative ayant un caractère de bilan, il est raisonnable de supposer qu'en principe seuls les objectifs qui sont le terme d'une unité d'enseignement et d'apprentissage devraient être soumis à l'évaluation sommative. Dans la même veine, les objectifs pré-requis

(logiquement ou psychologiquement) à l'atteinte de ces objectifs ultimes devraient être l'objet de l'évaluation formative. De plus, la non-maîtrise d'un objectif terminal pouvant nécessiter un enseignement correctif, on pourrait ajouter que certains objectifs terminaux seront également soumis à l'évaluation formative (avec des instruments distincts de ceux qui serviron à l'évaluation sommative pour ces objectifs). Enfin, certains objectifs intermédiaires qui, de réputation, ne présentent aucune difficulté particulière d'apprentissage seront soumis à une vérification de nature informelle (questions orales, simple exercice, observation directe etc.) alors que d'autres objectifs, plus difficiles à atteindre,feront l'objet d'une attention toute spéciale (instruments de nature critériée dont certains pourront prescrire un enseignement correctif particulier).

Voyons maintenant comment il est possible de représenter toutes ces nuances dans un seul tableau. La figure 6 (page 17) présente les 16 (seize) objectifs d'un cours disposés en réseaux ou hiérarchies. Dans ce diagramme les objectifs 2,7,12 et 16 sont des objectifs terminaux et seront visés par l'évaluation sommative (cercles contenant la lettre 'S'). Les objectifs intermédiaires 1, 3, 4, 5, 6, 9, 10, 11, 13, 14 et 15 feront l'objet de l'évaluation formative (cercles contenant la lettre 'F') (à noter l'abandon de l'objectif 8); certains de ces objectifs recevront un traitement spécial à cause de leur importance à être maîtrisés ou à cause de leur réputation de présenter des difficultés (cercles contenant le symbole 'F*' pour les objectifs 2, 6, 11 et 13).

Si l'évaluation sommative devait être réalisée à la fin de chaque étape d'une année scolaire et au moyen d'un examen de fin d'année, il est possible d'utiliser d'autres symboles pour désigner les objectifs qui feront l'objet de l'évaluation sommative en fin d'étape et les objectifs qui feront l'objet de l'évaluation sommative en fin d'année.

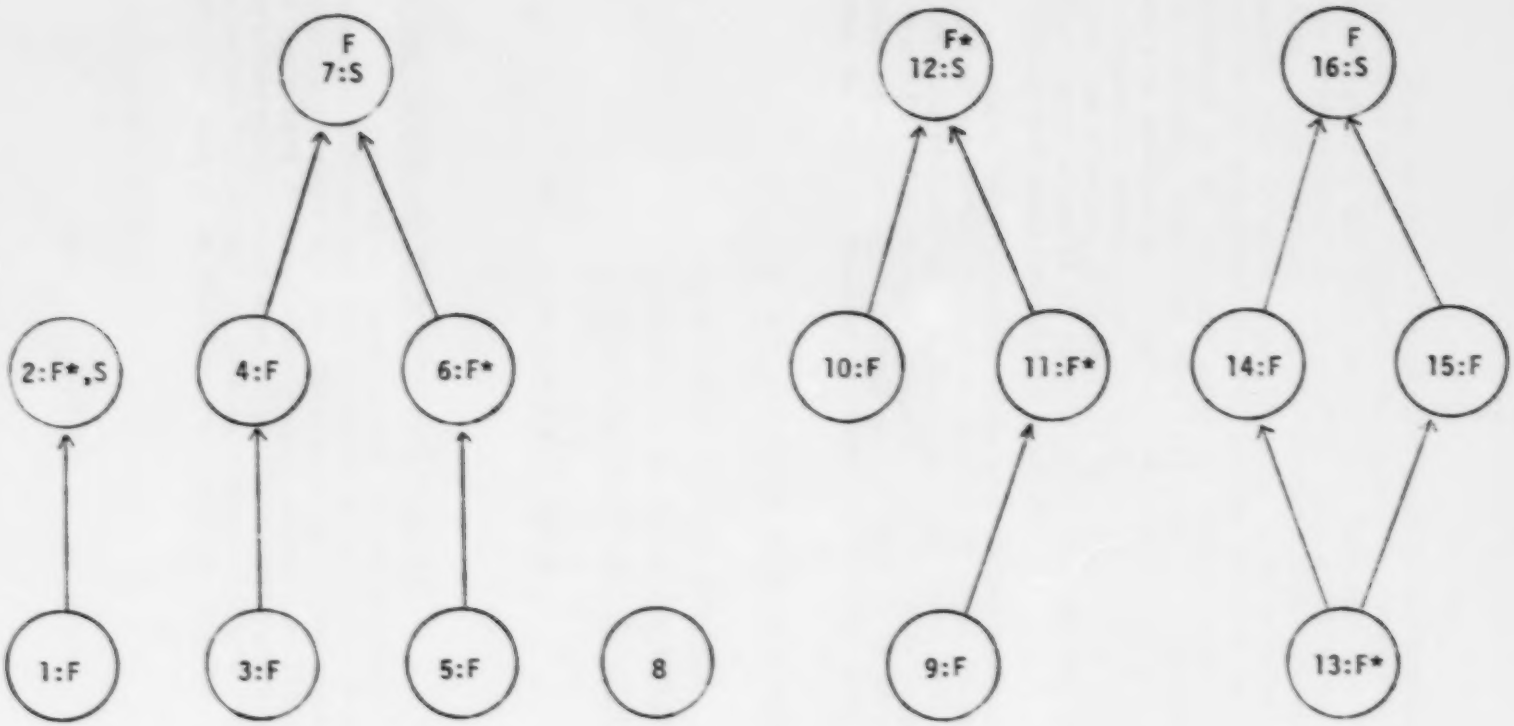


Figure 6: présentation des objectifs d'un cours en séquences d'objectifs intermédiaires et terminaux et détermination du type d'évaluation auquel ces objectifs seront soumis dans le cadre de l'évaluation des étudiants.

LES INSTRUMENTS DE L'EVALUATION FORMATIVE

Pour réaliser pleinement l'évaluation formative en milieu scolaire le degré d'atteinte des objectifs identifiés comme jalons d'un itinéraire doit être évalué avec soin pour que la décision de faire progresser ou non un étudiant vers d'autres objectifs soit valide. Dans plusieurs cas des procédés informels suffiront. Cependant, pour certains objectifs dont la maîtrise est jugée primordiale ou qui présentent des difficultés notoires au niveau de l'enseignement ou de l'apprentissage, il faudra recourir à des procédés plus raffinés. Nous avons recueilli dans la littérature un certain nombre de modèles qui présentent un intérêt particulier à l'égard du diagnostic pédagogique. Certains de ces modèles sont plus sophistiqués que d'autres et permettent de prescrire un enseignement correctif particulier là où la difficulté décelée chez un ou plusieurs étudiants est occasionnelle ou passagère.

Dans les prochaines pages nous allons exposer brièvement et illustrer un certain nombre de modèles de tests pouvant servir au diagnostic pédagogique, le choix que nous avons fait n'étant pas exhaustif. L'enseignement correctif que certains procédés pourraient prescrire est un sujet d'étude en lui-même et ne sera pas traité comme tel dans ce document. Les modèles retenus sont:

Le test centré sur un ou plusieurs objectifs,

La mesure reliée à un domaine,

L'analyse et la mesure des concepts,

La détection des erreurs caractéristiques,

et L'épreuve hiérarchique.

Le test centré sur un ou plusieurs objectifs.-

L'appellation "test centré sur les objectifs" traduit les expressions américaines "objective-based test" et "criterion-referenced test". A première vue tout instrument de mesure composé de questions se rapportant de façon congruente à quelques ou plusieurs objectifs opérationnels pourrait être qualifié de test "centré sur des objectifs". La nature "critériée" de tels instruments tient à l'interprétation du résultat observé en fonction d'un seuil de réussite. Il faut préciser

cependant que dans le cadre de l'évaluation formative l'interprétation critériée des résultats est une caractéristique nécessaire mais non suffisante du test centré sur un ou plusieurs objectifs (1).

En principe le test centré sur un ou plusieurs objectifs doit être conçu pour sonder (en profondeur autant que possible) le degré de maîtrise d'un ou de quelques objectifs d'apprentissage. Ce genre de test doit donc présenter idéalement les caractéristiques suivantes:

a) pour le test centré sur un seul objectif,

1- un certain nombre de questions congruentes à cet objectif,

2- la mention d'un seuil de réussite précisant le nombre minimum de questions (ou problèmes) que l'étudiant doit réussir pour affirmer qu'il a maîtrisé l'objectif: le verdict est qualitatif et signale le SUCCÈS ou l'ECHEC, la MAITRISE ou la NON-MAITRISE.

b) pour le test centré sur quelques objectifs,

1- un certain nombre de questions congruentes à chaque objectif et identifiées pour être groupées en sous-tests (dans l'épreuve même ou dans la grille diagnostique),

2- la mention d'un seuil de réussite pour chaque sous-test.

Il est important de souligner ici qu'avec plusieurs objectifs, c'est-à-dire avec plusieurs sous-tests, la correction des réponses de l'étudiant doit prévoir un résultat (score) pour chaque sous-test. Il faut comprendre qu'un score global aurait une faible valeur descriptive des difficultés que pourrait éventuellement éprouver un étudiant.

Le test centré sur un ou plusieurs objectifs présente certains avantages et aussi certaines limites. Ce type d'instrument est relativement facile à construire (en comparaison avec d'autres modèles pouvant être utilisés en évaluation formative). Il possède une faible valeur diagnostique car il ne permet pas à lui seul de prescrire un enseignement correctif qui soit autre que la reprise intégrale de l'apprentissage initial. Il peut cependant servir à vérifier le degré de maîtrise d'objectifs plus ou moins importants ou encore il peut s'avérer utile dans une démarche d'exploration diagnostique globale dans le but d'identifier le ou les objectifs qui mériteraient une attention particulière.

1- Peter W. Airasian et George F. Madaus, "Criterion-Referenced Testing in the Classroom", NCME Measurement in Education, Vol. 3, No. 4, 1972, pp. 3-4.

Du côté méthodologique la détermination du seuil idéal de réussite peut demander un travail considérable compte tenu des approches qui ont été suggérées dans la littérature. Plusieurs entreprises ou firmes spécialisées dans la production de tests critériés utilisent,comme politique, un seuil arbitraire qui peut varier entre 70% et 90% de réussites.

La page suivante contient un test de français centré sur trois objectifs et la grille diagnostique prévue pour la consignation des résultats de l'étudiant (à noter le verdict qualitatif qui est prévu pour chaque sous-test sur cette grille diagnostique).

La mesure reliée à un domaine (Domain-Referenced Testing).-

A l'origine (et encore aujourd'hui) la mesure reliée à un domaine prend tout son sens dans le concept d'univers d'items. Il est probable (car il serait difficile de l'affirmer catégoriquement) que les premiers concepteurs de ce type de mesure se soient préoccupés de l'attestation de compétence dans un contexte d'évaluation sommative. Le problème qui se pose alors est d'estimer le plus fidèlement possible la performance qu'un étudiant manifesterait s'il était théoriquement confronté à un univers d'items (ou domaine) et ce à partir d'un échantillon de cet univers. Dans ce contexte,la définition d'un domaine peut faire l'objet de spécifications très techniques dont le but est de créer une banque potentielle d'items en rapport avec un ou plusieurs objectifs. La littérature contient un assez bon nombre d'illustrations de domaines ou "d'item-forms".

Comme application possible au diagnostic pédagogique la mesure reliée à un domaine demeure très marginale puisque,jusqu'à ce jour,quelques chercheurs seulement ont tenté de semer l'idée (1)(2). Cette idée mérite d'être exploitée et nous allons tenter de l'exposer ici.

Le concept d'univers est fondamental pour saisir ce qu'est la mesure reliée à un domaine. Par exemple, l'ensemble de tous les problèmes d'addition qui pourraient être posés à un étudiant constitue un univers. L'ensemble de toutes les phrases simples que je peux imaginer pour demander à un étudiant d'identifier le mot ou les mots qui exercent la fonction "sujet du verbe" constitue un autre univers.

Il est possible d'effectuer une partition logique d'un univers en groupant les éléments qui ont une caractéristique commune. Dans le cas des problèmes d'addition, je peux séparer l'univers en deux sous-en-

---

1- Eva Lee Baker, "Beyond Objectives: Domain-Referenced Tests for Evaluation and Instructional Improvement", Educational Technology, 1974, pp. 10-16;

2- Louis Guttman et I.M. Schlesinger, The Analysis of Diagnostic Effectiveness of a Facet Design Battery of Achievement and Analytical Ability Tests, The Israel Institute of Applied Social Research, Jerusalem, Israel, 1967.

Objectifs mesurés par le test:

Parmi les mots d'une phrase simple, identifier le ou les mots qui sont...

01 sujet du verbe

02 complément d'objet direct

03 verbe.

Dans chacune des phrases suivantes, souligne le ou les mots qui sont sujet du verbe:

Pierre et Paul mangent des fruits.

La maison que nous avons achetée hier...

Bientôt le printemps diront les parents.

Dans chacune des phrases suivantes, souligne le complément d'objet direct:

J'ai bu de l'eau toute la soirée.

L'homme que tu as rencontré était poli.

Il a obtenu hier son augmentation.

Souligne le verbe dans les phrases suivantes:

Claude et moi irons au théâtre.

Les souris, disait-il, pourtant si délicates...

La veille de son départ il visita ses copains.

GRILLE DIAGNOSTIQUE étudiant(e)_____ [date]

Objectif	nombre de questions	seuil	résultat	[MAITRISE][✓]
identifier le sujet	3	3	____	[ ]
identifier le verbe	3	3	____	[ ]
identifier le c.o.d.	3	2	____	[ ]

sembles: les problèmes avec au moins une retenue et les problèmes sans retenue; dans le cas des phrases simples, je peux imaginer des phrases dans lesquelles le sujet est placé avant le verbe et d'autres phrases dans lesquelles le sujet est placé après.

Un même univers peut être partitionné de plusieurs façons et les partitions obtenues peuvent être combinées (croisées) les unes avec les autres. Le cas des phrases simples se prête à une telle combinaison. Par exemple, la sorte de mot pouvant être sujet du verbe (un nom, un pronom, plusieurs noms, etc.) est une base possible de partage de l'univers des phrases simples laquelle peut être combinée à la partition "sujet avant ou après le verbe". Le contenu de l'objectif suivant explicité delon deux dimensions (ou partages de l'univers) est un exemple de domaine (1):

Dans une phrase où il y a un ou plusieurs verbes, identifier le ou les mots qui sont sujets du verbe mis en évidence.

Remarque: selon le ou les mots employés et leur position dans la phrase, nous avons retenu plusieurs cas possibles pour décrire le sujet à identifier. Ces divers cas apparaissent dans le tableau suivant:

mots employés	position du sujet [2]	
	avant le verbe	après le verbe
nom seul		
pronom seul		
plusieurs noms		
plusieurs pronoms		
noms et pronoms		
infinitif		X X X [3]

1- cet exemple est extrait d'un document de l'Association des Institutions d'enseignement secondaire (A.I.E.S.), Diagnostic pédagogique et enseignement correctif, Document de sensibilisation, octobre 1979;

2- quelle que soit sa position, le sujet est toujours placé à proximité du verbe;

3- le cas de l'infinitif sujet placé après le verbe a été exclus du domaine.

En rédigeant plusieurs questions correspondant à chaque cellule du domaine (et congruentes avec l'objectif) on obtient un test centré sur un domaine, c'est-à-dire un ensemble de questions représentatif d'un univers quant aux dimensions qui ont été retenues. Vu sous cet angle ce type d'instrument serait un raffinement du test centré sur un objectif. Si les dimensions retenues présentent des degrés divers de difficulté les cellules du domaine donneront lieu à des sous-tests gradués en difficulté. Au mieux, au lieu de travailler avec les cellules du domaine, on peut partager les questions de l'épreuve entière en sous-tests en utilisant à tour de rôle chacune des dimensions. On voit donc que la conception d'un domaine à des fins diagnostiques présente beaucoup de possibilités. Un tel travail ne doit pas être laissé à la seule imagination du concepteur. Un enseignant expérimenté sera mieux inspiré que quiconque pour trouver des dimensions qui correspondent à une gradation réelle en difficulté.

Le test relié à un domaine doit être assorti d'une grille diagnostique pour consigner les résultats de l'étudiant. Le groupement des items en sous-tests permet de tracer un profil de l'étudiant pour chaque dimension. Les pages suivantes présentent une épreuve de 20 questions et la grille diagnostique correspondant à cette épreuve. L'épreuve a été rédigée à partir du domaine illustré précédemment avec le sujet du verbe.

L'interprétation des résultats consignés sur la grille diagnostique en rapport avec un test relié à un domaine souffre des problèmes qui sont inhérents à la détermination du seuil de réussite en mesure critériée. Toutefois la spécification du domaine selon des dimensions permet de sonder la ou les parties d'un univers d'items pour lesquelles l'étudiant peut éprouver des difficultés. Il est alors possible de prescrire à cet étudiant des exercices spécifiques en rapport avec ces difficultés. Ainsi, dans notre exemple, l'étudiant qui aurait échoué au sous-test "infinitif" (comme genre de mot pouvant être sujet) sera soumis à des exercices dans lesquels il devra identifier des "infinitifs sujets" d'une façon intensive. Enfin le choix de dimensions pertinentes pour circonscrire un domaine à des fins diagnostiques demande à être vérifié expérimentalement.

_____
(nom de l'élève)

Dans les phrases suivantes, écris, dans la colonne de droite, le ou les mots qui est(sont) sujet(s) du verbe souligné.

REMARQUE: pour que ta réponse soit jugée bonne, tu dois écrire seulement le ou les mots qui est(sont) sujet(s) sans en oublier aucun.

1.-	Il mange du spaghetti.	1)_____
2.-	La chaleur est étouffante pensent Marie, Jacques et lui.	2)_____
3.-	Travailler demande des efforts.	3)_____
4.-	A trois heures, les étudiants retournent à la maison.	4)_____
5.-	Ainsi marchent ceux et celles qui sont habitués aux longues excursions.	5)_____
6.-	Lui et toi allez à Québec.	6)_____
7.-	Manges-tu une pomme chaque jour?	7)_____
8.-	Pierre, Jean et Marie dégustent un sandwich à la cafétéria.	8)_____
9.-	Elles charment tous les jeunes gens par leur élégance.	9)_____
10.-	Pendant que flambent les maisons, les granges et les animaux, les habitants du village pleurent.	10)_____
11.-	Toi, lui et moi aimons danser à la discothèque.	11)_____
12.-	Paul, pense Jacques, est un garçon costaud.	12)_____
13.-	Dans la cour, René joue au hockey avec ses amis.	13)_____
14.-	Vus de haut, répondent elle et lui, les nuages sont très beaux.	14)_____
15.-	Qu'il pleuve pensent Marie, Nathalie et les autres!	15)_____
16.-	Josée, toi et moi marchons dans la forêt.	16)_____
17.-	Nager est agréable pendant la période estivale.	17)_____
18.-	Les écureuils et les hirondelles sont nombreux près du lac.	18)_____
19.-	Au pied de cette montagne coule une rivière.	19)_____
20.-	Ils sont fous ces Romains, s'exclame Astérix!	20)_____

Association des institutions d'enseignement secondaire [A.I.E.S.], tests diagnostiques: Français, secondaires I-II, octobre 1979.

GRILLE DIAGNOSTIQUE FR-AV-01

_____
(nom de l'élève)

SOUS-GROUPES DE QUESTIONS	nombre de questions	seuil de réussite	résultat obtenu	MAITRISE ✓
DIMENSION: mots qui sont sujets				
nom seul Nos: 4-12-13-19-20	5	5		
pronom seul Nos: 1-7-9	3	3		
plusieurs noms Nos: 8-10-18	3	2		
plusieurs pronoms Nos: 5-6-11-14	4	3		
nom(s) et pronom(s) Nos: 2-15-16	3	2		
infinitif Nos: 3-17	2	2		
DIMENSION: position du sujet par rapport au verbe				
le sujet est placé avant Nos: 1-3-4-6-8-9-11-13-16 17-18	11	10		
le sujet est placé après Nos: 2-5-7-10-12-14-15-19 20	9	7		

Association des institutions d'enseignement secondaire [A.I.E.S.], tests diagnostiques: Français, secondaires I-II, octobre 1979.

L'analyse et la mesure des concepts: le modèle de Markle et Tiemann(1).-

Un très grand nombre d'objectifs pédagogiques se situent au niveau du concept (ou notion). Ainsi on parlera du concept de complément d'objet direct, de nombre naturel, de masse, de démocratie, etc. La littérature qui a été publiée au sujet de l'apprentissage des concepts est des plus volumineuses. Il serait donc assez difficile de donner à un tel sujet un traitement équitable dans les limites du présent document. Nous allons nous en tenir à la démarche proposée par Markle et Tiemann pour analyser un concept à des fins d'enseignement et aussi à des fins d'évaluation. Sans entrer dans les détails nous présenterons dans les grandes lignes le fruit d'une telle démarche avec le concept d'adjectif qualificatif.

L'analyse du concept débute par l'explicitation des propriétés critiques (essentielles) et des propriétés non critiques (accessoires) qui font qu'une chose, un événement ou une personne constituent des exemples positifs ou des exemples négatifs du concept étudié. Dans le cas de l'adjectif qualificatif ce sont ces propriétés qui détermineront que dans certaines expressions le mot "ferme" est un adjectif qualificatif ou n'en est pas un.

Celui qui fait l'analyse du concept doit prévoir une banque d'exemples positifs et d'exemples négatifs qui serviront à l'enseignement et une autre banque d'exemples (positifs, négatifs) qui serviront à vérifier la maîtrise du concept par un étudiant. La méthodologie proposée par Markle et Tiemann se résume comme suit:

Analyse d'un concept: **********	PROPRIÉTÉS CRITIQUES
	PROPRIÉTÉS NON CRITIQUES
	EXEMPLES POSITIFS (enseignement)
	EXEMPLES NÉGATIFS (enseignement)
Mesure du concept: **********	Tâche d'identification contenant plusieurs exemples positifs et plusieurs exemples négatifs DIFFÉRENTS DE CEUX QUI ONT SERVI À L'ENSEIGNEMENT.

Si on applique cette méthodologie à l'analyse du concept d'adjectif qualificatif on obtient les éléments suivants (voir autre page):

1- Susan M. Markle et Philip W. Tiemann, Really Understanding Concepts: or in Frumious Pursuit of the Jabberwock, Tiemann Associates, Stipes Publishing Co., Champaign Illinois, 1970.

L'analyse du concept: **ADJECTIF QUALIFICATIF** (1)

PROPRIETES CRITIQUES	PROPRIETES NON CRITIQUES
(1) un mot	(4) accompagné ou non d'un déterminant
(2) qui se rapporte à un nom	(5) le nom est exprimé ou sous-entendu
(3) pour en exprimer une qualité bonne ou mauvaise ou une manière d'être	(6) placé avant ou après le nom (si ce dernier est exprimé

SOUS-ENSEMBLES D'EXEMPLES POSITIFS (fonction: faire varier chacune des prop. n.c.)

variation de (4)	Une abondante pluie Imprudents enfants qui osez...
variation de (5)	Il en a raconté de belles Il en a dit de bonnes choses Vous paraissez calmes
variation de (6)	Un magnifique jardin Un jardin magnifique

Sous-ENSEMBLES D'EXEMPLES NEGATIFS (fonction: exclure chacune des prop. c.)

absence de (1)	(exemples nég. exclus du domaine considéré ici)
absence de (2)	Parler haut Voir clair Plein les poches
absence de (3)	Cinq adolescents Par quelles villes avez-vous passé? Plusieurs personnes...
absence de (2) et (3)	(exemples nég. exclus du domaine considéré ici)

1- avec la collaboration de Claudette Gaudreault, étudiante à la Maîtrise en mesure et évaluation, Faculté des sciences de l'éducation de l'Université Laval, à l'hiver 1978.

Les exemples positifs et les exemples négatifs serviront à l'enseignement en attirant l'attention de l'étudiant sur le caractère essentiel ou accessoire de chaque propriété. La situation d'évaluation sera composée d'exemples distincts de ceux qui ont servi à l'enseignement (autrement on risquerait de ne mesurer que la "restitution" de réponses toutes faites dans le répertoire comportemental de l'étudiant). Voici un exemple de situation d'évaluation se rapportant à l'adjectif qualificatif:

> Coche entre parenthèses ( ) la ou les expressions dans lesquelles le mot souligné est un adjectif qualificatif.
>
> (N'INDIQUE RIEN EN CAS DE DOUTE)
>
> 1.- Quelques personnes m'ont dit... ( )
> 2.- Ambitieux, ces enfants iront loin. ( )
> 3.- J'en ai vu de plus terribles. ( )
> 4.- Chanter juste. ( )
> 5.- Un discours intéressant. ( )
> 6.- Cette vue est fort belle. ( )
> 7.- L'hôpital contenait quatorze lits. ( )
> 8.- Un ancien hôtel. ( )
> 9.- Je le crois honnête. ( )

La page suivante présente le corrigé et certaines prévisions des erreurs que pourrait commettre un étudiant qui ne maîtrise pas le concept d'adjectif qualificatif. Selon le type d'erreur manifestée, l'étudiant sera exposé à des exemples nouveaux (positifs ou négatifs) qui serviront à attirer son attention sur des propriétés qui lui ont échappé au cours de l'apprentissage initial.

Le problème majeur que pose l'utilisation du modèle de Markle et Tiemann pour le diagnostic pédagogique est que les correctifs à apporter sont tributaires du développement théorique de l'apprentissage des concepts. Il existe en effet plusieurs stratégies d'enseignement susceptibles de prévenir ou de contrer certains types d'erreurs conceptuelles. Il est probable également que le modèle de Markle et Tiemann soit limité à un mode particulier d'apprentissage de concepts: l'apprentissage par induction. Enfin, dans la pratique, le comportement de

certains élèves à une épreuve construite selon ce modèle a révélé que les réussites et les échecs ne permettaient pas toujours de dire avec certitude si ces élèves étaient caractérisés par des erreurs de surgénéralisation ou des erreurs de sousgénéralisation. La mesure des concepts à des fins de diagnostic pédagogique demeure un sujet de recherche des plus intéressants.

Corrigé et interprétation diagnostique

1.- Quelques personnes...	( )	absence de la prop. c. #3
2.- Ambitieux, ces enfants..	(✓)	
3.- ..de plus terribles.	(✓)	
4.- Chanter juste	( )	absence de la prop. c. #2
5.- ..discours intéressant	(✓)	
6.- ...fort belle	( )	absence de la prop. c. #2
7.- ...quatorze lits	( )	absence de la prop. c. #3
8.- Un ancien hôtel	(✓)	
9.- Je le crois honnête.	(✓)	

ERREURS DE SOUS GENERALISATION:

omission de choisir 3 et 9 (ét. exige que le nom soit exprimé)

omission de choisir 2 et 8 (ét. exige que l'ad. qual. soit placé après le nom)

autres... ???

ERREURS DE SURGENERALISATION:

choix de 4 et 6 (ét. néglige prop. c. #2--'se rapporte à un nom')

choix de 1 et 7 (ét. néglige prop. c. #3--'qualité ou manière d'être)

autres... ???

CONFUSION: combinaison des deux types d'erreurs.

La détection des erreurs caractéristiques: la rédaction systématique des leurres pour des questions à choix de réponse.-

Il existe plusieurs situations où des étudiants, lorsque confrontés à un type de problème, ont tendance à commettre la même erreur et ce d'une façon constante. Le cas le plus typique est celui de la multiplication de termes algébriques où il a été observé que certains étudiants ont la fâcheuse habitude de multiplier les exposants au lieu de les additionner.

L'idée des erreurs que peut commettre de façon caractéristique un étudiant n'est pas nouvelle. On la retrouve dans des procédés de correction de résolution de problèmes (ou de réponses élaborées à des questions) qui accordent un crédit partiel au processus de solution utilisé par l'étudiant au lieu de ne corriger que la réponse uniquement. Avec un seul problème posé à plusieurs étudiants on peut déceler une ou plusieurs erreurs communes à ces étudiants. Dans un tel cas on ne saurait parler d'erreur caractéristique pour un étudiant en particulier.

Pour déceler une ou des erreurs caractéristiques que pourrait éventuellement manifester un étudiant il faut présenter à celui-ci un certain nombre de problèmes du même type. La correction du procédé de solution utilisé par l'étudiant pour résoudre chacun des problèmes peut révéler une tendance de la part de celui-ci à répéter le même genre d'erreur. Ce genre d'erreur sera dit caractéristique de cet étudiant.

Le problème méthodologique qui se pose alors est de déterminer le nombre minimal de fois (fréquence minimale), pour un ensemble de questions ou problèmes d'un même type, à partir duquel on peut inférer qu'il s'agit d'une erreur caractéristique. Ce problème est analogue à celui que pose la détermination d'un seuil de réussite en mesure critériée.

Dans certains cas, avec des problèmes ou des questions d'un même type, il est possible d'anticiper des solutions erronées (ce qu'un enseignant expérimenté peut faire) et de prévoir ainsi des réponses qui trahissent la solution utilisée par l'étudiant. On pourra alors utiliser la forme "question à choix de réponse", forme de question au moyen de laquelle on propose à l'étudiant:

a) une bonne réponse placée au hasard parmi les réponses suggérées et ce différemment d'une question à l'autre,

b) une mauvaise réponse pour chaque solution erronée anticipée (ces mauvaises réponses sont également disposées au hasard),

c) un choix "aucune de ces réponses" dans l'éventualité où un étudiant commettrait une combinaison d'erreurs ou toute erreur non anticipée.

La forme "question à choix de réponse" pour détecter des erreurs caractéristiques a été suggérée par Guttman et Schlesinger(1). L'épreuve diagnostique rédigée selon ce procédé est de correction facile. L'inconvénient réside dans le fait qu'elle ne permet pas de déceler des combinaisons d'erreurs. Dans certains cas il sera recommandé de recourir aux formes de question "à réponse brève" ou "à réponse élaborée" qui se prêtent à une correction qui peut être fastidieuse.

La page suivante présente un test diagnostique qui a été construit selon un procédé de rédaction systématique des leurres. Sur l'autre page on présente un problème type ainsi que la description des solutions anticipées pour rédiger les leurres de chaque problème du test, ainsi que la bonne réponse. En dessous du problème type apparaît la grille diagnostique qui serait utilisée pour consigner les choix de l'étudiant avec des commentaires décrivant le type d'erreur.

Le problème type qui a été utilisé pour illustrer un test diagnostique construit avec des leurres systématiques est une adaptation d'un exemple suggéré par Davis et Diamond (2).

L'épreuve hiérarchique.-

L'épreuve hiérarchique, comme son nom l'indique, est un test dont les questions se rapportent aux objectifs d'une hiérarchie d'apprentissages. Il s'agit essentiellement d'un test centré sur plusieurs objectifs et l'interprétation diagnostique de la copie d'un étudiant réside dans la réussite aux sous-tests. Cependant l'épreuve hiérarchique va plus loin en ce sens qu'elle permet de déterminer le niveau auquel se situe un étudiant dans une hiérarchie d'apprentissages, à la condition, bien sûr, que l'épreuve ou plus précisément la hiérarchie soit valide.

Les problèmes méthodologiques rattachés à l'épreuve hiérarchique sont nombreux. Nous n'en signalerons que quelques uns:

1- il y a le problème de l'établissement de la hiérarchie d'objectifs à partir d'un objectif terminal,

2- pour chaque élément de la hiérarchie il faut déterminer le nombre minimal de questions et le seuil minimal de performance afin de tirer des conclusions valides,

3- la validation même de la hiérarchie: exhaustivité et liens hiérarchiques entre les objectifs.

1- Louis Guttman et I.M. Schlesinger, "Systematic Construction of Distractors for Ability and Achievement Test Items", Educational and Psychological Measurement, Vol. XXVII, 1967, pp. 569-580;

2- Frederick B. Davis et James J. Diamond, "The Preparation of Criterion-Referenced Tests", pp. 116-138 dans: C.W. Harris, M.C. Alkin et W.J. Popham (eds), Problems in Criterion-Referenced Measurement, Los Angeles: UCLA Graduate School of Education, Center for the Study of Evaluation, 1974.

****
* Pour chacun des problèmes suivants encerclez la lettre *
* qui correspond à votre réponse. *
****

(1) Un poste récepteur de radio coûte $ 50.00. Le vendeur exige un accompte de $ 10.00 et quatre versements mensuels de $ 12.50 chacun. Le prix initial de l'appareil est augmenté de quel pourcentage avec cette manière de payer?

a) 0 % ou 0.00
b) 16 2/3 % ou 0.16 2/3
c) 20 % ou 0.20
d) 500 % ou 5.00
e) autre réponse:

_____
(écrivez votre réponse ici)

(2) Pour se procurer une montre il a fallu payer un montant de $ 28.00 par mois pendant trois mois après avoir versé une somme de $ 16.00 au moment de l'achat. Sachant que la montre coûtait initialement $ 80.00, quel est le pourcentage d'augmentation du prix de la montre qui correspond à ce mode de paiement?

a) 400 % ou 4.00
b) 25 % ou 0.25
c) 20 % ou 0.20
d) 5 % ou 0.05
e) autre réponse:

_____
(écrivez votre réponse ici)

(3) Un tourne-disques valant $ 100.00 a été acheté par une personne qui a versé $ 10.00 en accompte et 13 versements de $ 10.00; cette façon de payer a augmenté de quel pourcentage le prix du tourne-disques?

a) 40 % ou 0.40
b) 30 $ ou 0.30
c) 28 4/7 % ou 0.28 4/7
d) 250 % ou 2.50
e) autre réponse:

_____
(écrivez votre réponse ici)

(4) Pour obtenir une collection de livres on a le choix entre deux modes de paiement:

Mode A: payer comptant une somme de $ 120.00

Mode B: payer un accompte de $ 20.00 et faire 14 versements mensuels consécutifs de $ 10.00 chacun.

Le prix de la collection est augmenté de quel pourcentage avec le mode B ?

a) 300 % ou 3.00
b) 16 2/3 % ou 0.16 2/3
c) 25 % ou 0.25
d) 33 1/3 % ou 0.33 1/3
e) autre réponse:

_____
(écrivez votre rép. ici)

Exemple de problème: Le prix d'un fauteuil est $ 60; un client dépose un accompte de $ 9 et s'engage à verser par la suite un montant de $ 7 par mois et ce pendant 9 mois. Cette façon de payer augmente le coût du fauteuil de quel pourcentage?

Liste des solutions possibles: BONNE RÉPONSE et ERREURS CARACTÉRISTIQUES

I	solution correcte	$\frac{9 + 7 \times 9 - 60}{60}$ x 100:	20 %
II	oublie l'accompte	$\frac{7 \times 9 - 60}{60}$ x 100:	5 %
III	% du total payé au lieu du prix	$\frac{9 + 7 \times 9 - 60}{72}$ x 100:	16 2/3 %
IV	Rapport en I inversé	$\frac{60}{9 + 7 \times 9 - 60}$ x 100:	500 %
V	autres réponses: erreur de calcul et/ou combinaison d'erreurs		---

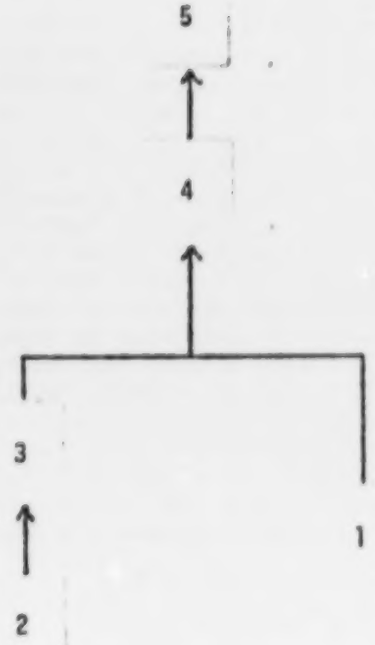
CHOIX DE L'ETUDIANT (1)	(2)	(3)	(4)	TOTAL	COMMENTAIRES
c	b	a	d	____	(solutions correctes)
a	d	b	b	____	(accompte oublié !)
b	c	c	c	____	(% du total payé au lieu du prix)
d	a	d	a	____	(notion de % d'une quantité ???)
e	e	e	e	____	(erreur de calcul ou autres types d'erreur)

Si l'objectif cible du test diagnostic est un objectif terminal, le test hiérarchique sera particulièrement utile pour prescrire un enseignement correctif au niveau de certains objectifs intermédiaires.

Nous ne fournirons pas d'exemple de test hiérarchique. En se servant de la hiérarchie qui est illustrée à la figure 5 (page 14) le lecteur n'a qu'à imaginer un test qui serait composé d'environ 3 ou 4 questions par objectif. La grille diagnostique qui accompagnerait un tel test contiendrait en énumération tous les patrons de SUCCES et d'ECHECS qui sont compatibles avec cette hiérarchie. Ces patrons déterminent d'ailleurs divers niveaux de réussite des objectifs de la hiérarchie. Pour un étudiant donné il suffirait de noter le patron de réussites-échecs qui le caractérise.

Le diagramme ci-dessous présente un exemple de grille diagnostique qui pourrait accompagner un test hiérarchique construit en fonction de la hiérarchie présentée à la figure 5.

nom de l'étudiant: _____



Interprétation diagnostique:

NIVEAUX D'ATTEINTE DES OBJECTIFS DE LA SEQUENCE PROPOSEE

(si la hiérarchie est valide!)

objectif				
1	2	3	4	5
E	E	E	E	E
S	E	E	E	E
E	S	E	E	E
S	S	E	E	E
S	S	S	E	E
S	S	S	S	E
S	S	S	S	S

S: atteinte de l'objectif
E: non atteinte de l'objectif

L'UTILISATION DES RÉSULTATS EN ÉVALUATION FORMATIVE: la feuille de route individuelle et la matrice "étudiants x objectifs".

Nous avons vu précédemment que pour interpréter un échec dans l'atteinte d'un objectif et en tirer une déduction hypothétique valable il nous faut deux types d'informations:

1- le rendement de l'étudiant à d'autres objectifs afin de déterminer si cet échec est la manifestation d'une déficience généralisée chez cet étudiant ou s'il s'agit d'une difficulté passagère,

2- le rendement d'autres étudiants face à l'objectif d'intérêt afin de déterminer si la difficulté éprouvée par l'étudiant se présente comme un cas isolé (individuel) ou fait partie d'un phénomène collectif.

Les rapports d'évaluation permettant d'avoir accès à ces deux types d'informations sont la feuille de route individuelle de l'étudiant et la matrice "étudiants x objectifs". Il existe malheureusement peu d'exemples de rapports d'évaluation dans la littérature qui répondent adéquatement à ce besoin d'informations. Les rapports d'évaluation suggérés par les firmes spécialisées qui produisent des tests critériés constituent des modèles intéressants à exploiter dans le cadre de l'évaluation formative des étudiants. Certaines formules devront cependant être adaptées car il s'agit habituellement de tests du type "survey" qui mesurent à l'occasion d'une opération globale un très grand nombre d'objectifs à la fois, ce qui n'est pas à proprement parler dans l'esprit de l'évaluation formative. De plus, certaines firmes comme Science Research Associates et CTB McGraw-Hill offrent des services de correction et de codification aux usagers de leurs tests et produisent à cette occasion des rapports statistiques assez détaillés et susceptibles de satisfaire les besoins de plusieurs institutions. Le diagnostic, qui est une facette de cette production, revêt un aspect de bilan et vise surtout à donner une image assez fidèle des objectifs maîtrisés par un ou plusieurs groupes d'étudiants avec la possibilité de comparer les résultats obtenus dans une institution à ceux obtenus dans d'autres institutions.

Compte tenu de la nature et de l'esprit même de l'évaluation formative il serait souhaitable que les informations à recueillir au sujet de la maîtrise ou de la non-maîtrise d'objectifs soient notées sur un rapport d'évaluation à chacun des jalons d'un itinéraire donné et ce d'une manière cumulative.

La page suivante présente un exemple de feuille de route pouvant être utilisée dans le cadre d'une séquence d'apprentissages de base en arithmétique. Les objectifs de la séquence sont énumérés avec concision et un symbole quelconque (un crochet comme dans l'exemple) indique les objectifs maîtrisés par l'étudiant à un moment donné. Ce modèle de feuille de route individuelle est réduit à sa plus simple expression et les symboles appropriés sont ajoutés, c'est-à-dire que les objectifs maîtrisés sont notés, à mesure que l'étudiant progresse.

Il est possible de raffiner la feuille de route de plusieurs façons. Dans un système d'enseignement individualisé on pourrait mentionner la date à laquelle l'étudiant a maîtrisé chaque objectif mettant ainsi en relief le rythme de croisière de l'étudiant dans sa progression. Il est possible également (et il sera parfois utile) d'indiquer sur la feuille de route, en regard de chaque objectif mentionné, le nombre d'essais qui ont été requis à l'étudiant pour maîtriser l'objectif. Il suffit d'utiliser à cet effet une symbolique qui permette de différencier l'étudiant qui réussit du premier coup le test de maîtrise de l'étudiant qui réussit le test de maîtrise (ou une forme équivalente de ce test) après avoir été soumis à une ou plusieurs séances d'enseignement correctif. Ce genre d'information pourrait présenter l'aspect suivant:

Etudiant:_____

objectif:	1er test	2ième test	3ième test
.....	S		
.....	S		
.....	E	S	
.....	S		
.....	E	E	S

etc. ...

S: succès E: échec (ou non maîtrise)

EXEMPLE DE FEUILLE DE ROUTE

apprentissages de base en arithmétique

Nom de l'élève: _____ Date: _____

Compter des objets jusqu'à 5 . . . . . ✓

Choisir un nombre d'objets jusqu'à 10. . . . . ✓

Lire et écrire les nombres jusqu'à 50. . . . . ✓

Identifier les symboles de l'addition et de la soustraction . . . . . ✓

Compter des objets...

par groupes de 2 jusqu'à 20 . . . . . _____

par groupes de 10 jusqu'à 100 . . . . . ✓

Reconnaître les pièces de monnaie (1¢ 5¢ 10¢ 25¢). . ✓

Faire des combinaisons d'addition (jeu de 10) . . . _____

Faire des combinaisons de soustraction (jeu de 10) . _____

Donner le moment de la journée...

à l'heure près . . . . . ✓

à la demi-heure près . . . . . _____

au quart d'heure près . . . . . _____

Démontrer, avec des objets concrets, la valeur des unités et des dizaines dans un nombre compris entre 1 et 50 . . . . . _____

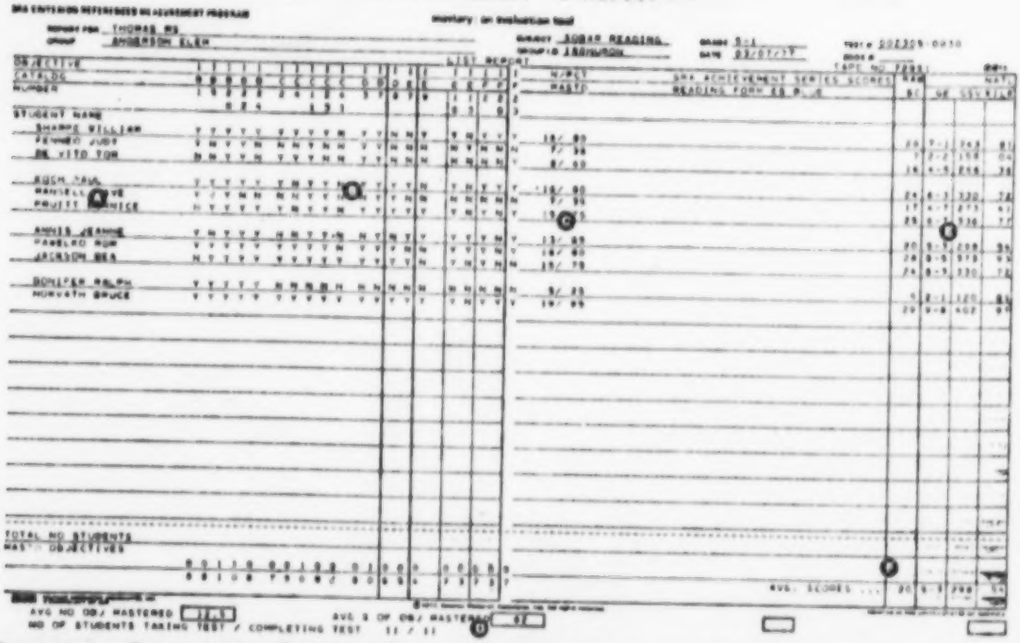
Ecrire les chiffres romains jusqu'à XII . . . . . _____

Il faut comprendre que le but poursuivi en notant de telles informations n'est pas d'étiqueter l'étudiant mais de justifier l'interprétation qui pourrait être faite d'un échec à un objectif au moment où cet échec se produit. L'étudiant qui aurait maîtrisé les objectifs antérieurs mais au prix de plusieurs efforts devrait recevoir une attention particulière qui va au-delà du correctif pédagogique en rapport avec un seul objectif non maîtrisé (à moins que les efforts déployés par l'étudiant,comme l'indique sa feuille de route,soient un phénomène collectif: ce qui amènerait une remise en question de l'organisation des objectifs et/ou de la pédagogie).

La matrice "étudiants x objectifs", qui elle aussi peut être complétée de façon cumulative, indiquera pour chaque étudiant une séquence de réussites et d'échecs aux objectifs d'un cours ou d'une séquence d'apprentissage. Si tout le groupe d'étudiants est considéré, on aura le nombre ou le pourcentage d'étudiants ayant maîtrisé chaque objectif, indiquant ainsi les objectifs pour lesquels l'échec ou la non maîtrise est soit un phénomène collectif soit le triste sort d'un seul ou de quelques étudiants. Dans un système d'enseignement individualisé, où chaque étudiant progresse à son propre rythme, l'information de groupe pour un objectif donné ne sera pas nécessairement disponible au moment d'interpréter un échec en particulier. Il faudra alors avoir recours à des instruments de mesure qui ont été utilisés plus d'une fois et pour lesquels on connaît le taux de réussite dans un groupe repère.

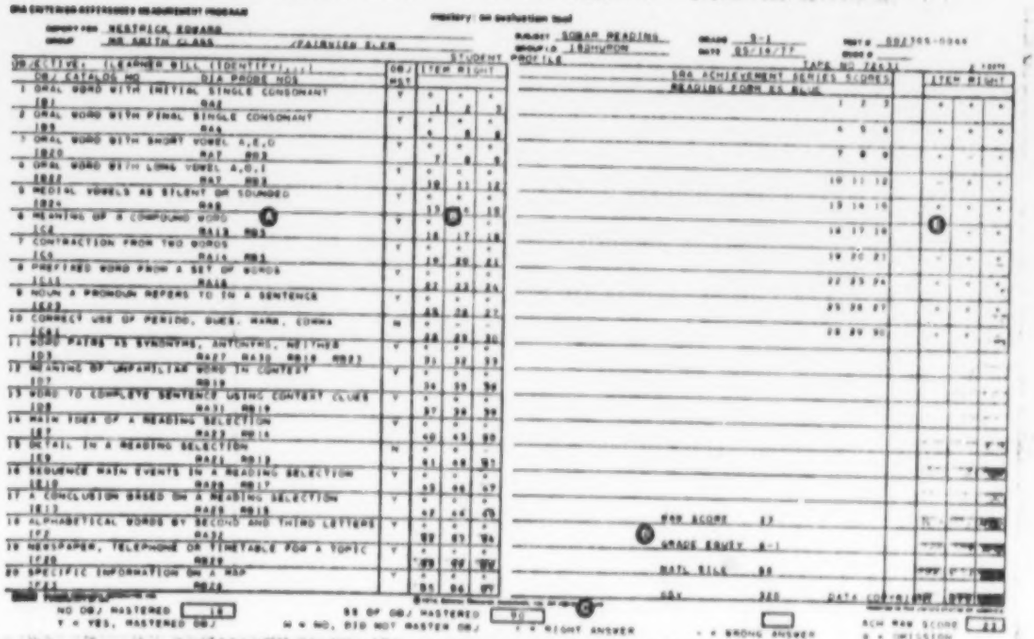
La page suivante est un extrait de la documentation fournie par Sciences Research Associates, Inc. (catalogue 1978). Le tableau du haut est un exemple de matrice "étudiants x objectifs" sur laquelle les succès (Y pour yes) et les échecs (N pour no) ont été notés. Le tableau du bas présente un profil individuel un peu plus sophistiqué que la feuille de route donnée en exemple à la page précédente car la réussite à chaque question du test centré sur les objectifs y est notée.

NRT/CRT List Report



Ⓐ STUDENT NAME Ⓑ STUDENT PERFORMANCE IN TERMS OF CRT OBJECTIVES Ⓒ STUDENT PERFORMANCE ON OVERALL CRT Ⓓ GROUP PERFORMANCE ON OVERALL CRT Ⓔ STUDENT PERFORMANCE ON OVERALL NRT Ⓕ GROUP PERFORMANCE ON OVERALL CRT

NRT/CRT Individual Student Profile



Ⓐ LIST OF OBJECTIVES FOR CRT Ⓑ STUDENT PERFORMANCE ON CRT ITEMS AND OBJECTIVES Ⓒ OVERALL STUDENT PERFORMANCE ON CRT Ⓓ OVERALL STUDENT PERFORMANCE ON NRT Ⓔ STUDENT PERFORMANCE ON NRT ITEMS